QR Factorization (or encoding Gram-Schmidt)

A QR factorization of an $m \times n$ matrix A with linearly independent columns is:

$A = QR$, where Q has orthonormal columns, R is invertible upper triangular

(A: $m \times n$, Q: $m \times n$, R: $n \times n$)

This is obtained via Gram-Schmidt.

Let $A = [a_1 \cdots a_n]$

Apply Gram-Schmidt to $\{a_1, \ldots, a_n\}$ to get $\{q_1, \ldots, q_n\}$.

Let $Q = [q_1 \cdots q_n]$.

Since $\{q_1, \ldots, q_n\}$ is an orthonormal basis of W and $a_j \in W$,

$$a_j = \text{proj}_W a_j = \text{proj}_{q_1} a_j + \cdots + \text{proj}_{q_n} a_j$$
$$= (a_j \cdot q_1) q_1 + \cdots + (a_j \cdot q_n) q_n$$

By construction $q_i \cdot a_j = 0$ for all $i > j$. So:

$$a_j = (a_j \cdot q_1) q_1 + \cdots + (a_j \cdot q_j) q_j$$

Set $R = \begin{bmatrix} q_1 \cdot a_1 & q_1 \cdot a_2 & \cdots & q_1 \cdot a_n \\ 0 & q_2 \cdot a_2 & \cdots & q_2 \cdot a_n \\ \vdots & & \ddots & \vdots \\ 0 & - & - & q_n \cdot a_n \end{bmatrix}$

(i,j)-th entry is $q_i \cdot a_j$

Then:

$$QR = Q \begin{bmatrix} q_1 \cdot a_1 & q_1 \cdot a_2 & \cdots & q_1 \cdot a_n \\ 0 & q_2 \cdot a_2 & \cdots & q_2 \cdot a_n \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & q_n \cdot a_n \end{bmatrix}$$

The i^{th} column of QR is

$$Q \begin{bmatrix} q_1 \cdot a_i \\ \vdots \\ q_i \cdot a_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} = (q_1 \cdot a_i) q_1 + (q_2 \cdot a_i) q_2 + \cdots + (q_i \cdot a_i) q_i = a_i \; !$$

Hence $A = QR$.

Note! To find R, simply compute $Q^T A$,

since $Q^T A = Q^T (QR) = (Q^T Q) R = I \cdot R = R$.

Ex: $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. By the previous example,

Gram-Schmidt yields $Q = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/2 \\ 0 & 1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & 1/2 \\ 0 & 1/\sqrt{2} & -1/2 \end{bmatrix}$

Hence $R = Q^T A = \begin{bmatrix} \sqrt{2} & \sqrt{2} & -\sqrt{2}/2 \\ 0 & \sqrt{2} & 3\sqrt{2}/2 \\ 0 & 0 & 1 \end{bmatrix}$ (Check)

Why do this? It makes least squares problems easier to solve.

Say $A = QR$.

Then to find the least squares solution to $Ax = b$,

we need to solve $A^TA\hat{x} = A^Tb$. Let's simplify this expression:

$$A^TA\hat{x} = A^Tb$$
$$(QR)^T(QR)\hat{x} = (QR)^Tb$$
$$(R^TQ^T)(QR)\hat{x} = R^TQ^Tb$$
$$R^T(Q^TQ)R\hat{x} = R^TQ^Tb$$
$$R^TIR\hat{x} = R^TQ^Tb$$
$$R^TR\hat{x} = R^TQ^Tb$$
$$(R^T)^{-1}R^TR\hat{x} = (R^T)^{-1}R^TQ^Tb$$
$$R\hat{x} = Q^Tb$$

So we need only solve $R\hat{x} = Q^Tb$

Solving $R\hat{x} = Q^Tb$ is extremely simple since R is upper triangular. Just use back substitution!

(this is what computers do)

Ex: Find a least square solution to:

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

We know $Q = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/2 \\ 0 & 1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & 1/2 \\ 0 & 1/\sqrt{2} & -1/2 \end{bmatrix}$

and $R = \begin{bmatrix} \sqrt{2} & \sqrt{2} & -\sqrt{2}/2 \\ 0 & \sqrt{2} & 3\sqrt{2}/2 \\ 0 & 0 & 1 \end{bmatrix}$

Since $Q^T b = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/2 \end{bmatrix}$,

we need to solve $\begin{bmatrix} \sqrt{2} & \sqrt{2} & -\sqrt{2}/2 \\ 0 & \sqrt{2} & 3\sqrt{2}/2 \\ 0 & 0 & 1 \end{bmatrix} \hat{x} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/2 \end{bmatrix}$

which can be done via backward substitution.

Determinants

Recall if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det A = ad - bc$.

Let's interpret this quantity geometrically.

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by $T(v) = Av$.

Let $C = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \mid 0 \le x, y \le 1 \right\} = \left\{ x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mid 0 \le x, y \le 1 \right\}$ (unit square)

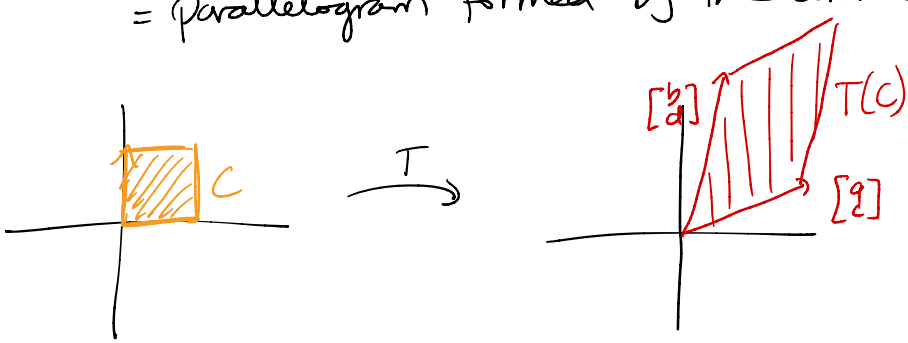
Then $T(C) = \left\{ T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) \mid 0 \le x, y \le 1 \right\}$

$= \left\{ T\left(x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \mid 0 \le x, y \le 1 \right\}$

$= \left\{ xT\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + yT\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \mid 0 \le x, y \le 1 \right\}$ since T is linear

$= \left\{ x\begin{bmatrix} a \\ c \end{bmatrix} + y\begin{bmatrix} b \\ d \end{bmatrix} \mid 0 \le x, y \le 1 \right\}$

= parallelogram formed by the columns of A.

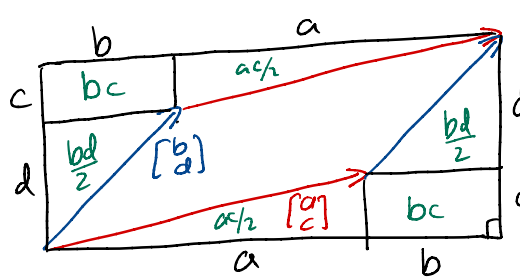


C has area 1. What is the area of $T(C)$?

Assume $a, b, c, d > 0$. Then the area of $T(C)$ is

$(a+b)(c+d) - 2bc - ac - bd = ad - bc$

$= \det A$!



More generally, area of $T(C) = |\det A|$.